Exam 2 will cover sections 1.1-1.5, 2.1-2.9, 3.1, 3.2, 4.1 and 4.2 through Factorization theorem (through p. 115), with emphasis on sections 2.6, 2.7, 2.9. 3.1, 3.2, 4.1 and 4.2 through p. 115. Steiner's formula and iterated expectations will not be on the exam.

No notes will be allowed, but bring a calculator. You should memorize the pmf or pdf f, $E(Y)$ and $\mathrm{V}(Y)$ for the following RVs: 1) beta(δ, ν),
2) Bernoulli(ρ) = bin$(k = 1, \rho)$, 3) binomial(k, ρ), 4) Cauchy(μ, σ),
5) chi-square(p) = gamma$(\nu = p/2, \lambda = 2)$, 6) exponential(λ) = gamma$(\nu = 1, \lambda)$,
7) gamma(ν, λ), 8) $N(\mu, \sigma^2)$, 9) Poisson(θ), and 10) uniform(θ_1, θ_1).

You should memorize the mgf of the binomial, χ^2_p, exponential, gamma, normal and Poisson distributions. You should memorize the cdf of the exponential and of the normal distribution $\Phi(\frac{y-\mu}{\sigma})$.

Types of problems:

1) If $Y_1, ..., Y_n$ are independent with mgfs $m_{Y_i}(t)$, then the mgf of $W = \sum_{i=1}^n Y_i$ is $m_W(t) = \prod_{i=1}^n m_{Y_i}(t)$. See HW4 1, Q4.

2) If $Y_1, ..., Y_n$ are iid with mgf $m_Y(t)$, then the mgf of $W = \sum_{i=1}^n Y_i$ is $m_W(t) = [m_Y(t)]^n$. The mgf of $\overline{Y}$ is $m_{\overline{Y}}(t) = [m_Y(t/n)]^n$. See HW4 2,3,4, Q4.

3) Given $m_Z(t)$, if $m_Z(t)$ is a "brand name mgf", you should be able to give the distribution of Z. For example, if $m_Z(t) = \exp(5t + 50t^2)$, then Z is normal with mean $\mu = 5$ and variance $\sigma^2 = 100$.

4) Suppose $W = \sum_{i=1}^n Y_i$ or $W = \overline{Y}$ where $Y_1, ..., Y_n$ are independent. For several distributions (especially Y_i iid gamma(ν, λ) and Y_i independent $N(\mu_i, \sigma_i^2)$), you should be able to find the distribution of W, the mgf of W, $E(W)$, Var(W), and $E(W^2) = V(W) + [E(W)]^2$. See Q4 3.

5) If $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and if $\boldsymbol{A}$ is a $q \times p$ matrix, then $\boldsymbol{AX} \sim N_q(\boldsymbol{A\mu}, \boldsymbol{A\Sigma A}^T)$. If $\boldsymbol{a}$ is a $p \times 1$ vector of constants, then $\boldsymbol{a} + \boldsymbol{X} \sim N_p(\boldsymbol{a} + \boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Suppose

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \ and \ \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

6) $\boldsymbol{X}_1 \sim N_q(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$.

7) Given a MVN distribution, be able to find the MVN distribution of subsets, pairs of independent random variables and the correlation
$\rho(X_i, X_j) = \frac{\sigma_{i,j}}{\sqrt{\sigma_{ii}\sigma_{jj}}} = Cov(X_i, X_j)/\sqrt{V(X_i)V(X_j)}$. See HW4 5.

8) If $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then the conditional distribution of $\boldsymbol{X}_1$ given that $\boldsymbol{X}_2 = \boldsymbol{x}_2$ is multivariate normal with mean $\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)$ and covariance $\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. That is,

$$\boldsymbol{X}_1 | \boldsymbol{X}_2 = \boldsymbol{x}_2 \sim N_q(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}).$$

Usually $p = 2$. Often $X_1 = Y$ and $X_2 = X$. Then find $E(Y|X), V(Y|X)$ and $\rho(Y, X)$. See HW4 6 and Q4 1.

9) Know that (X,Y) can have a joint distribution that is not multivariate normal, yet the marginal distributions of X and Y are both univariate normal. Hence X and Y can be normal, but $aX+bY$ is not normal. (Need the joint distribution of (X,Y) to be MVN for all linear combinations to be univariate normal.) See Remark 2.1.

10) Know that if $Y_1,...,Y_n$ are iid with $E(Y)=\mu$ and $\text{V}(Y)=\sigma^2$, then $E(\overline{Y})=\mu$ and $\text{V}(\overline{Y})=\sigma^2/n$. Know $E(S^2)=\sigma^2$.

11) See Q5, HW5 1-9. Given the pmf or pdf of Y or that Y is a brand name distribution, know how to show whether Y belongs to an exponential family or not using $f(y|\boldsymbol{\theta}) = h(y)c(\boldsymbol{\theta})\exp[\sum_{i=1}^k w_i(\boldsymbol{\theta})t_i(y)]$ or $f(y|\theta)=h(y)c(\theta)\exp[w(\theta)t(y)]$.

Tips: a) The F, Cauchy, logistic, t and uniform distributions can not be put in form 11) and so are not exponential families.

b) If the support depends on an unknown parameter, then the family is not an exponential family.

12) If Y belongs to an exponential family, know how to find the natural parameterization, $\eta_i = w_i(\boldsymbol{\theta})$ and the natural parameter space Ω. Find Ω by finding the ranges of η_i. For a kP-REF, Ω is an open set that is typically a cross product of k open intervals. For a 1P-REF, Ω is an open interval. See HW5 1-9, Q5.

13) If Y belongs to an exponential family, know how to show whether Y belongs to a k-dimensional regular exponential family. In particular, if $p=2$ you should be able to show whether η_1 and η_2 satisfy a linearity constraint and whether t_1 and t_2 satisfy a linearity constraint, to plot Ω and to determine whether the natural parameter space contains a 2-dimensional rectangle.

Tip: If the family is a 2 dimensional exponential family with natural parameters η_1,η_2 but the usual parameterization is determined by 1 parameter θ, then the family is probably not regular. $N(\mu,\mu^2)$ is a common example. If $dim(\Theta)=j<k=dim(\Omega)$, the family is usually not regular. See HW5 1-9, Q5.

14) If $Y_1,...,Y_n$ are iid $N(\mu,\sigma^2)$ know that $\overline{Y} \sim N(\mu,\sigma^2/n)$ and $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$. Then $S^2 \sim \sigma^2\chi^2_{n-1}/(n-1)$ and $\sum_{i=1}^n (Y_i-\overline{Y})^2 \sim \sigma^2\chi^2_{n-1}$. See HW5 10.

15) If $Y_1,...,Y_n$ are iid, then $f_{Y_{(n)}}(t) = n[F_Y(t)]^{n-1}f_Y(t)$, $F_{Y_{(n)}}(t)=[F_Y(t)]^n$, $f_{Y_{(1)}}(t) = n[1-F_Y(t)]^{n-1}f_Y(t)$ and $F_{Y_{(1)}}(t) = 1-[1-F_Y(t)]^n$. You may need to find $E(Y_{(n)})$ or $E(Y_{(1)})$. See Q5 and the example done in class.

16) Know what a sufficient statistic is, and know how to use the Factorization theorem $f(\boldsymbol{y}|\boldsymbol{\theta}) = g(\boldsymbol{T}(\boldsymbol{y})|\boldsymbol{\theta})\ h(\boldsymbol{y})$ to find a sufficient statistic $\boldsymbol{T}(\boldsymbol{Y})$ for $\boldsymbol{\theta}$ if $Y_1,...,Y_n$ are iid with pdf or pmf $f_Y(y)$. See HW5 11, Q5.

Tips: i) for iid data with marginal support $\mathcal{Y}_i \equiv \mathcal{Y}^*$, $I_{\mathcal{Y}}(\boldsymbol{y}) = I(\text{all } y_i \in \mathcal{Y}^*)$. If $\mathcal{Y}^* = (a,b)$, then $I_{\mathcal{Y}}(\boldsymbol{y}) = I(a<y_{(1)}<y_{(n)}<b) = I(a<y_{(1)})I(y_{(n)}<b)$. Put $I(a<y_{(1)})$ in $g(\boldsymbol{T}(\boldsymbol{y})|\boldsymbol{\theta})$ if a is an unknown parameter but put $I(a<y_{(1)})$ in $h(\boldsymbol{y})$ if a is known. If both a and b are unknown parameters, put $I(a<y_{(1)}<y_{(n)}<b)$ in $g(\boldsymbol{T}(\boldsymbol{y})|\boldsymbol{\theta})$. If $b=\infty$, then $I_{\mathcal{Y}}(\boldsymbol{y}) = I(a<y_{(1)})$. If $\mathcal{Y}^*=[a,b]$, then $I_{\mathcal{Y}}(\boldsymbol{y}) = I(a \le y_{(1)} < y_{(n)} \le b) = I(a \le y_{(1)})I(y_{(n)} \le b)$. ii) Try to make the dimension of $\boldsymbol{T}(\boldsymbol{y})$ as small as possible. Put anything that depends on $\boldsymbol{y}$ but not $\boldsymbol{\theta}$ into $h(\boldsymbol{y})$.